

A Presumed DNA Helicase Encoded by *ERCC-3* Is Involved in the Human Repair Disorders Xeroderma Pigmentosum and Cockayne's Syndrome

Geert Weeda,*† Reinier C. A. van Ham,*
Wim Vermeulen,‡ Dirk Bootsma,‡
Alex J. van der Eb,* and Jan H. J. Hoeijmakers‡

* Laboratory for Molecular Carcinogenesis
Medical Genetics Center
Sylvius Laboratory
P. O. Box 9503
2300 RA Leiden
The Netherlands

‡ Department of Cell Biology and Genetics
Medical Genetics Center
Erasmus University Rotterdam
P. O. Box 1738
3000 DR Rotterdam
The Netherlands

Summary

The human gene *ERCC-3* specifically corrects the defect in an early step of the DNA excision repair pathway of UV-sensitive rodent mutants of complementation group 3. The predicted 782 amino acid *ERCC-3* protein harbors putative nucleotide, chromatin, and helix-turn-helix DNA binding domains and seven consecutive motifs conserved between two superfamilies of DNA and RNA helicases, strongly suggesting that it is a DNA repair helicase. *ERCC-3*-deficient rodent mutants phenotypically resemble the human repair syndrome xeroderma pigmentosum (XP). *ERCC-3* specifically corrects the excision defect in one of the eight XP complementation groups, XP-B. The sole XP-B patient presents an exceptional conjunction of two rare repair disorders: XP and Cockayne's syndrome. This patient's DNA contains a C→A transversion in the splice acceptor sequence of the last intron of the only *ERCC-3* allele that is detectably expressed, leading to a 4 bp insertion in the mRNA and an inactivating frameshift in the C-terminus of the protein. Because XP is associated with predisposition to skin cancer, *ERCC-3* can be considered a tumor-preventing gene.

Introduction

An intricate network of biochemical repair systems operates in cells to protect DNA from deleterious accumulation of DNA damage and permanent mutations. One of the most important and best studied repair processes in prokaryotes and eukaryotes is the nucleotide excision repair pathway (reviewed by Friedberg, 1985). This system removes a broad range of lesions caused by very dissimilar agents such as UV-induced photoproducts (cyclobutane and 6–4 pyrimidine dimers, thymine glycols), bulky chemical adducts, and cross-links. It is thought to sense

general distortions in the regular helical conformation of DNA (e.g., kinks and bends). A detailed insight into the molecular mechanism of this process has emerged from the work on *Escherichia coli* (reviewed by Sancar and Sancar, 1988; Grossman et al., 1988). In this organism at least six proteins participate in this pathway, including the SOS-induced *uvrAB* helicase complex that scans the DNA for structural abnormalities and the *uvrC* endonuclease that incises the damaged strand on both sides of the lesions.

Little is known about the way the nucleotide excision machinery acts in eukaryotes. In the yeast *Saccharomyces cerevisiae*, mutant analysis and gene cloning have revealed the existence of at least ten genetic loci collectively designated as the *RAD3* epistasis group (Friedberg, 1988; Haynes and Kunz, 1981). Two classes of excision repair-defective cell lines can be discerned within mammals. The first comprises cells from patients suffering from the autosomal-recessive human repair disorder xeroderma pigmentosum (XP). This rare, cancer-prone syndrome is clinically characterized by extreme sensitivity of the skin to sunlight (UV) exposure, pigmentation abnormalities, predisposition to skin cancer, and, frequently, neurological degeneration (Cleaver and Kraemer, 1989). The molecular defect in most XP patients resides in early steps of the excision repair pathway. Genetic complementation analysis by cell fusion has revealed the existence of eight complementation groups, designated XP-A through XP-H. (Recently, the status of complementation group XP-H has been challenged [Johnson et al., 1989; Robbins, 1989; Johnson, 1989; our unpublished data].) Interestingly, two patients belonging to distinct XP complementation groups, XP-B and XP-H, have been found to exhibit the clinical symptoms of both XP and another rare excision repair disorder, Cockayne's syndrome (CS) (reviewed by Lehmann, 1987).

The second class of mammalian repair mutants is represented by the laboratory-induced UV-sensitive rodent (mainly Chinese hamster ovary [CHO]) cell lines, among which eight complementation groups have been thus far identified (reviewed by Collins and Johnson, 1987; Busch et al., 1989). It is unlikely that with these sets of mutant cells all the mammalian repair functions are represented. Complementation tests between the rodent and XP mutants—although far from complete—have thus far not revealed any overlap between these two classes of repair-deficient cells (Thompson et al., 1985; Stefanini et al., 1985; van Duin et al., 1989). These findings suggest the involvement of a substantial number of genes in excision repair and a considerable complexity at the biochemical level.

The rodent mutants have proven to be very valuable for the isolation of correcting human genes, designated excision repair cross-complementing rodent repair deficiency (*ERCC*) genes. DNA-mediated gene transfer has led to the cloning of the *ERCC-1* and *ERCC-2* genes, correcting CHO mutants belonging to complementation groups 1

† Present address: Department of Cell Biology and Genetics, Erasmus University, 3000 DR Rotterdam, The Netherlands.

and 2, respectively (Westerveld et al., 1984; Weber et al., 1988). Analysis of these genes has revealed a striking evolutionary sequence conservation at the protein level. The *ERCC-1* gene product is homologous to the yeast *RAD10* repair protein and to parts of the *E. coli* *uvrA* and *uvrC* polypeptides (van Duin et al., 1986; Hoeijmakers et al., 1986; Doolittle et al., 1986). The *ERCC-2* gene was found to be the human counterpart of the yeast repair gene *RAD3* (Weber et al., 1990), which encodes a 5'→3' DNA helicase (Sung et al., 1987a, 1987b). In addition, the mouse gene correcting SV40-transformed XP-A fibroblasts has been cloned after extensive genomic DNA transfection experiments (Tanaka et al., 1989). The cloning of other XP genes is being attempted (Arrand et al., 1989; Teitz et al., 1987). Recently, we have cloned the *ERCC-3* gene, correcting CHO mutants belonging to group 3 (Weeda et al., 1990). Here we report the cDNA cloning, sequence analysis, and involvement of the *ERCC-3* gene in XP and CS.

Results

Cloning of *ERCC-3* cDNA

The human excision repair gene *ERCC-3* was cloned after DNA-mediated gene transfer of HeLa chromosomal DNA into the UV-sensitive CHO mutant 27-1, a member of complementation group 3 (generously provided by Dr. R. D. Wood, ICRF, London) (Weeda et al., 1990). Mutants of this group are very sensitive to UV irradiation and to agents inducing bulky adducts; they have reduced DNA repair synthesis and harbor a defect in one of the steps preceding the incision of damaged DNA. These characteristics strongly resemble those of cells from XP patients.

Unique *ERCC-3* genomic sequences were used as probes to isolate the *ERCC-3* cDNA from various human cDNA libraries (see Experimental Procedures and Weeda et al., 1990, for details). The physical maps of the four cDNA clones discussed in the present article are depicted in Figure 1. The insert size of pCD1 (~2.8 kb) is very close to the size of *ERCC-3* mRNA (2.9 kb), suggesting that this cDNA is (close to) full-length. Furthermore, transfection of *ERCC-3* cDNA clones pCD1 and pH3T or pHK1 inserted in mammalian expression vectors conferred wild-type UV survival and repair synthesis to 27-1 cells with a very high efficiency (Weeda et al., 1990). Plasmids pCD2 and pH3T1, with shorter inserts, were unable to induce a repair-proficient phenotype, suggesting that they are incomplete.

Nucleotide Sequence of *ERCC-3* cDNA

The nucleotide sequence and the deduced amino acid sequence of pCD1 and pH1 are shown in Figure 2. The first ATG (position 40) starts an open reading frame (ORF) of only 48 bp, whereas the second ATG (position 96) is followed by an ORF of 2346 bp. Although in more than 95% of all mRNAs known the first ATG serves as the start codon for translation (Kozak, 1984), we believe that in the case of *ERCC-3* the second ATG represents the actual translation initiation site of the *ERCC-3* protein. This proposition is based on the following observations. First,

the purine residue found at the -3 position of most eukaryotic start codons (Kozak, 1984) is not present in front of the first ATG (in contrast to the second ATG; Figure 2A). In other examples where the 5' ATG is in an unfavorable context for initiation (i.e., position -3 not A/G, position +4 not G), "leaky scanning" has been observed, which enables some ribosomes to reach and initiate at the more favorable second ATG (Kozak, 1989). Second, a data base search of protein coding regions based on codon preference (Staden and McLachlan, 1982) gave a strong bias in favor of the second, long ORF. Third, the first ATG is absent from the 5' leader of the functional mouse *ERCC-3* cDNA (unpublished data). Finally, we found a shorter human cDNA clone (pHK1) that lacks the first ATG but still confers wild-type UV resistance to 27-1 cells. Notably, the human *ERCC-2* gene harbors ATG codons in its 5' untranslated region as well (Weber et al., 1990). Inspection of the 3' region revealed the sequence of the common polyadenylation signal (Wickens and Stephenson, 1984) at 29 bp before the start of the poly(A) tail. The pentanucleotide sequence CAYTG, which is found adjacent to the polyadenylation site in many eukaryotic genes (Berg, 1984), is also present. A group of AU-rich sequences (AUUUA) in the 3' noncoding region of many short-lived mRNAs associated with metabolic instability (Shaw and Kamen, 1986) can be detected in the 3' noncoding region of *ERCC-3* (Figure 2A).

Sequence of *ERCC-3* Protein and Homology with Other Proteins and Functional Domains

The deduced amino acid sequence of the *ERCC-3* cDNA predicts a protein of 782 amino acids, with a calculated molecular mass of 89,274 kd. To determine whether the *ERCC-3* gene product shows sequence similarity with other polypeptides, a FASTA computer search was performed. No strongly homologous proteins were found with published sequences available in various data bases. A systematic search for functional domains present in the predicted amino acid sequence of *ERCC-3* yielded the following regions of interest:

First, there is a potential nuclear localization signal at amino acid positions 6–18. A stretch of basic amino acids has been identified in several proteins involved in mediating transport to the nucleus (see Roberts, 1989, for a recent review). In Figure 2C the amino acid sequence of part of the *ERCC-3* protein is compared with a number of known and presumed nuclear localization signals in other proteins. The motif KK/RXK/R has emerged as a common (but not exclusive) feature in many nuclear localization sequences described to date and is found in the exposed N-terminal part of *ERCC-3*.

Second, there are several regions with a pronounced negative charge (amino acids 20–29, 256–265, 697–701, and 721–728). Stretches of acidic amino acids have been found in a number of proteins that associate with chromatin and/or histones such as the high mobility group I family (Wen et al., 1989), transcription factors (for a review see Ptashne, 1988), nucleolin (Lapeyre et al., 1987), and the histone 2A/2B-specific ubiquitin-conjugating enzyme *RAD6* (Jentsch et al., 1987). The negatively charged amino

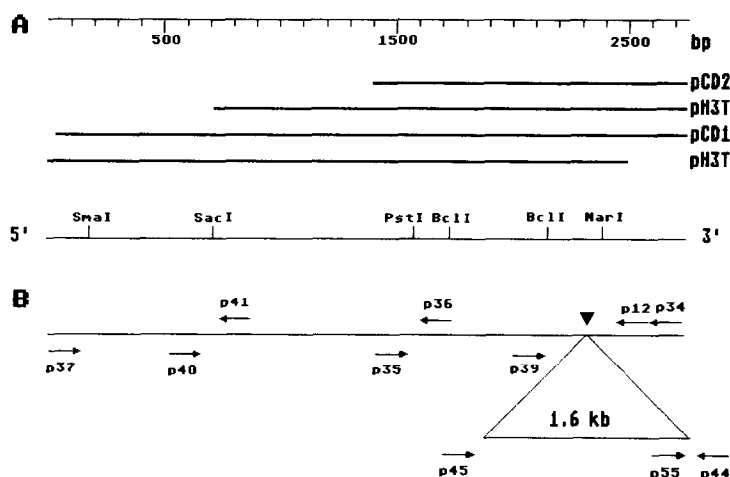


Figure 1. *ERCC-3* cDNA

(A) Screening of human cDNA libraries with genomic or cDNA *ERCC-3* probes yielded five overlapping clones: pCD1, pCD2, pH3T, pH3T1, and pHK1 (the last is not shown, since it is comparable to pCD1 apart from lacking 71 bp at the 5' terminus). Restriction enzyme maps for the insert region are diagrammed. Clones pCD1 and pHK1 were sequenced in both orientations. Except for pCD1 and pCD2, cDNA inserts were cloned into the eukaryotic expression vector pSVL. The ability of the cDNAs to correct the UV sensitivity of 27-1 cells to wild-type resistance by transfection with the cDNA expression clones is indicated with + or -. cDNA clones pHK1, pH3T, and pH3T1 are flanked by *EcoRI* linkers.

(B) Strategy for amplification of *ERCC-3* cDNA and gene sequences. Sequences of the oligonucleotide primers appear in Experimental Procedures. Oligonucleotide primers p34, p37, and p39 contain a restriction site for *HindIII*, *EcoRI*, and *BamHI*, respectively. The 4 bp insertion is indicated with a triangle.

acids are thought to permit electrostatic interaction with the basic histones. The presence of such modules in *ERCC-3* suggests that it has the capacity to associate with chromatin.

Third, there is a potential DNA binding motif at amino acids 121–142. In Figure 2D the amino acid sequence of this part is compared with a number of known or presumed “helix-turn-helix” DNA binding domains. This motif is one of the most extensively investigated DNA binding domains identified in a number of prokaryotic and eukaryotic proteins, including prokaryotic repressors, yeast mating type regulatory proteins, and homeobox proteins (for reviews see Pabo and Sauer, 1984; Struhl, 1989). *ERCC-3* contains the correct amino acids at the position considered to be crucial for the configuration of the two helices, consistent with the idea that this part may be involved in DNA binding.

Fourth, inspection of the *ERCC-3* main ORF revealed the “A” site of the NTP binding sequence typical of numerous enzymes utilizing ATP and GTP (Walker et al., 1982). Alignment with the nucleotide binding site of the yeast ssDNA-dependent ATPase *RAD3*, which possesses DNA helicase activity, and with the consensus sequence deduced from many ATPases is shown in Figure 2E, domain I. Furthermore, a segment corresponding to the “B” site of NTP binding domains, which is thought to interact with the Mg^{2+} cation of Mg-NTP through the conserved D residue (de Vos et al., 1988), is readily recognized in the *ERCC-3* amino acid sequence (segment II in Figure 2E).

Fifth, the “A” nucleotide binding site starts with a region exhibiting a virtually perfect match with seven consecutive domains recently recognized by several investigators to be conserved between two superfamilies of known and putative DNA and RNA helicases (Hodgman, 1988; Gorbelenya et al., 1988, 1989; Linder et al., 1989). The deduced consensus sequence of each module and the cor-

responding sequence of the yeast repair helicase *RAD3* are compared with that of *ERCC-3* in Figure 2E. It is evident that *ERCC-3* matches overall better to the consensus than *RAD3*. Some *ERCC-3* domains display—apart from the consensus residues—extensive additional similarity to *RAD3* (e.g., domains I, IA, III, and particularly VI, of which 10 of the 17 residues are identical). The remaining portions of the two proteins are, however, very different. The presence of all motifs, their proper positioning, and their high level of homology to the consensus sequence provide strong evidence for a potential DNA unwinding function of *ERCC-3*.

These data taken together suggest that the *ERCC-3* gene encodes a nuclear, DNA and chromatin binding helicase, with a well-organized overall structure; its central part starts with the nucleotide binding domain, encoding a DNA unwinding function, which is preceded by a DNA binding motif and a nuclear localization sequence, and which is flanked by acidic regions that enable the protein to interact with (basic) histones (Figure 2B). However, definite proof for these properties awaits functional analysis of the protein.

Role of *ERCC-3* in XP

Since the repair phenotype of rodent group 3 mutants resembles that of XP cells in many aspects, it was of interest to examine whether *ERCC-3* is involved in this human disorder. To answer that question, a systematic analysis was performed at the DNA, RNA, and functional levels. First, we searched for gross alterations (deletions and rearrangements detectable by Southern blot analysis) of the *ERCC-3* gene in representative cell lines of all excision-deficient XP complementation groups. Figure 3A shows that *EcoRI* digests of chromosomal DNA of the XP cells displayed the same hybridization pattern with an *ERCC-3* cDNA probe as a digest of DNA from HeLa control

A

5'
 888agcttccggattgagccggagatccccccagagcgagatgcccggcgccgctgtggagcggggtcatcttctctctgtctgttagtgcacATGGGCAAAAGACACCGACCGGACC 120
 NLS
 R D K K K S R K R H Y E D E E D D E E D A P G N D P Q E A V P S A A G K Q V D E 48
 GCGACAAGAGAAATCCAGCAAGCGGCACTATGAGGATGAAGAGGATGATGAAGAGGACGCCCGGGGAACGACCTCAGGAAGCGGTTCCTCGGCGGCGGGGAAGCAGGTGGATGAGT 240
 S G T K V D E Y G A K D Y R L Q M P L K D D H T S R P L W V A P D G H I F L E A 88
 CACGCCAAAGTGGATGAATATGGAGCCAAGGACTACAGGCTGCAAAATGCCGTGAAGGACGACACACCTCCAGGCCCTCTGGTGGCTCCCGATGGCCATATCTTCTTGAAGCCT 360
 DNA-BD
 F S P V Y K Y A Q D E L V A I A E P V C R P T H V H E Y K L T A Y S L Y A A V S 128
 TCTCTCCAGCTTACAAATATGCCAAGACTTCTTGGTGGCTATTGACAGCCAGTGTGCGGACCAACCATGTGCATGAGTACAACTAAGTCTACTCTTGTATGCAGCTGTACGG 480
 V G L Q T S D I T E Y L R K L S K T G V P D G I M Q F I K L C T V S Y G K V K L 168
 TTGGCTGCAAAACAGTACATCCGAGTACCTCAGCAAGCTCAGCAAGCTGAGTCCCTCATGGAATATGCAGTTTATTAAGTCTGTACTGTACGTATGGAAGAACTCAAGCTGG 600
 V L K H N R Y F V E S C H P D V I Q H L L Q D P V I R E C R L R N S E G E A T E 208
 TCTGAAGCACAAGATCTCTGTTAAAGTTGCCACCTGATGAATCCAGCATCTCTCCAGGACCCGTGATCCGGAATGCCCTTAAGAACTCTGAAGGGGAGGCCATGAGC 720
 L I T E T F T S K S A I S K T A E S S G G P S T S R V T D P Q G K S D I P M D L 248
 TCATCAGAGACTTTCAAGCAAAATCTGCCATTTCTAAGACTGCTGAAGCAGTGGTGGCCCTCCACTTCCCGAGTGACAGATCCACAGGTAAATCTGACATCCCATGACCTGT 840
 F D F Y E Q M D K D E E E E E E E T Q T V S F E V K Q E M I E E L Q K R C I H L E 288
 TTGACTTCTATGACAAATCGACAAGATGAAGAAGAAAGAGACACAGACAGTGTCTTTTGAAGTCAAGCAGGAATGATTGAGGAATCCAGAACTGTGATCCACCTGGAGT 960
 Y P L L A E Y D F R N D S V N P D I N I D L K P T A V L R P Y Q E K S L R K M F 328
 ACCCTCTGTGCGAGAATATGACTTCGGGAATGATTCTGTCAACCTGATATCAACATTGACCTAAAGCCCAAGCTGTCTCAGACCTATCAGGAGAAGAGCTTCGCAAGATGTTCT 1080
 G N G R A R S G V I V L P C G A G K S L V G V T A A C T V R K R C L V L G N S A 368
 GAAACGGGCGTGACGTTCCGGGCTATTGTTCTTCCCTGCGGTGCTGGAAGTCCCTGGTGTGTGACTGCTGCATGCAGTGTACAGAAACGCTGTCTGTGCTGGGCAACTCAGCTG 1200
 V S V E Q W K A Q F K M W S T I D D S Q I C R F T S D A K D K P I G C S V A I S 408
 TTTCTGTGGAGCTGGAAGCCCAATTCAGATGTGGTCCACCATGACGACAGCCAGATCTGCGCGTTCACTCCGATGCCAAGGACAAGCCCATCGGCTGCTCCGTGTCATAGCA 1320
 T Y S M L G H T T K R S W E A E R V M E W L K T Q E W G L M I L D E V H T I P A 448
 CCTACTCATGCTGGCCACACCAACCAAGGTCTCTGGAGGCGGAGCGAGTCTGAGTGGCTCAAGACCCAGGATGGGCGCTCATGCTCGATGAAGTGCACACCATACCA 1440
 K M F R R V L T I V Q A H C K L G L T A T L V R E D D K I V D L N F L I G P K L 488
 AGATGTTCCGAAGGCTGCTCAACATGCTGAGGCCCACTGTAAAGTGGGTTGACTGGGACCTCGTCCGGAAGATGACAAAATTTGGATTAAATTTTCTGATTGGGCTAAGCTCT 1560
 Y E A N W M E L Q N N G Y I A K V Q C A E V W C P M S P E F Y R E Y V A I K T K 528
 ACGAAGCCAATCGATGAGCTGCAAGATATGGCTACATCGCCAAAGTCCAGTGTGCTGAGGCTGTGTCGCCATGTCTCTGAAATTTACGGGAATATGTGGCAATCAAAACCAAGA 1680
 K R I L L Y T M N P N K F R A C Q F L I K F H E R R N D R I I V F A D N V F A L 568
 AACGAATCTGCTGTACACCATGAACCCCAACAAATTTAGAGCTTCCAGTTTCTGATCAAGTTTCATGAAGGAGGAATGACAAGATTATTGTCTTGTGCAATGTGTTGCCCTAA 1800
 K E Y A I R L N K P Y I Y G P T S Q G E R M Q I L Q N F K H N P K I N T I F I S 608
 AGGAATATGCCATTCCACTGAACAAACCTATATCTACGGACCTACGCTCTCAGGGGGAAGATGCAAAATCTCCAGAATTTCAGGACAAACCCCAAAATTAACACCATCTCATATCA 1920
 K V G D T S F D L P E A N V L I Q I S S H G C S R R R Q E A Q R L G R V L R A K K 648
 AGGTAGGTGACACTTCGTTGATCTGCGGAAGCAATGTCTCTCAGATCTCATCCATGCTGCTCCAGGCGTCAGGAAGCCCAAGGCTAGGGCGGGTGTCTCGAGCTAAAAAG 2040
 G M V A E E Y N A F F Y S L V S Q D T Q E M A Y S T K R Q R F L V D Q G Y S F K 688
 CGATGTTTCAGAGAGTACAATCCCTTTTCTACTCACTGGTATCCAGGACACACAGGAATGGCTTACTCAACCAAGGGCAGAGATTCTTGTAGATCAAGGTTATAGCTTCAAG 2160
 V I T K L A G M E E E D L A F S T K E E Q Q Q L L Q K V L A A T D L D A E E E V 728
 TGATCAACAACTCCCTGGCTGAGGAGCAAGACTTGGCTTTTTCGACAAAAGCAAGCAACAGCAGCTCTTACAGAAAGTCTTGGACGCCACTGACCTGGATGCCAGGAGAGCTGG 2280
 V A G E F G S R S S Q A S R R F G T H S S M S G A D D T V Y M E Y H S S R S K A 768
 TGGCTGGGAATTTGGCTCCAGATCCAGCCAGGATCTGGCCCTTTGGACCATGAGTTCTATGCTCTGGGCGGACGACACTGTACATGGAGTACCCTCATCGCGGACCAAGCGCG 2400
 P S K H V H P L F K R P R K * 782
 CCAGCAAAATCTACACCCGCTTCTCAAGCGCTTATGAAATGATgcttagggcgaggtacttcttcaagaccggcgcttggcacccttcttggaaagggtatttcagcataacatttcc 2520
 cttccaccttttgccttccctccagcgttggccaaattgtgtgaggaagatcatcaaggcttggctgtgcttccataggtcatctagggtttttataaggaggaggagacaatat 2640
 tttttcaacttttttggggagtggttcttctgtatataaaaaattgttaatttttaagggtattttatgttaaccgttctgtgataaacaggaatggacccttgaccagct 2750
 3'

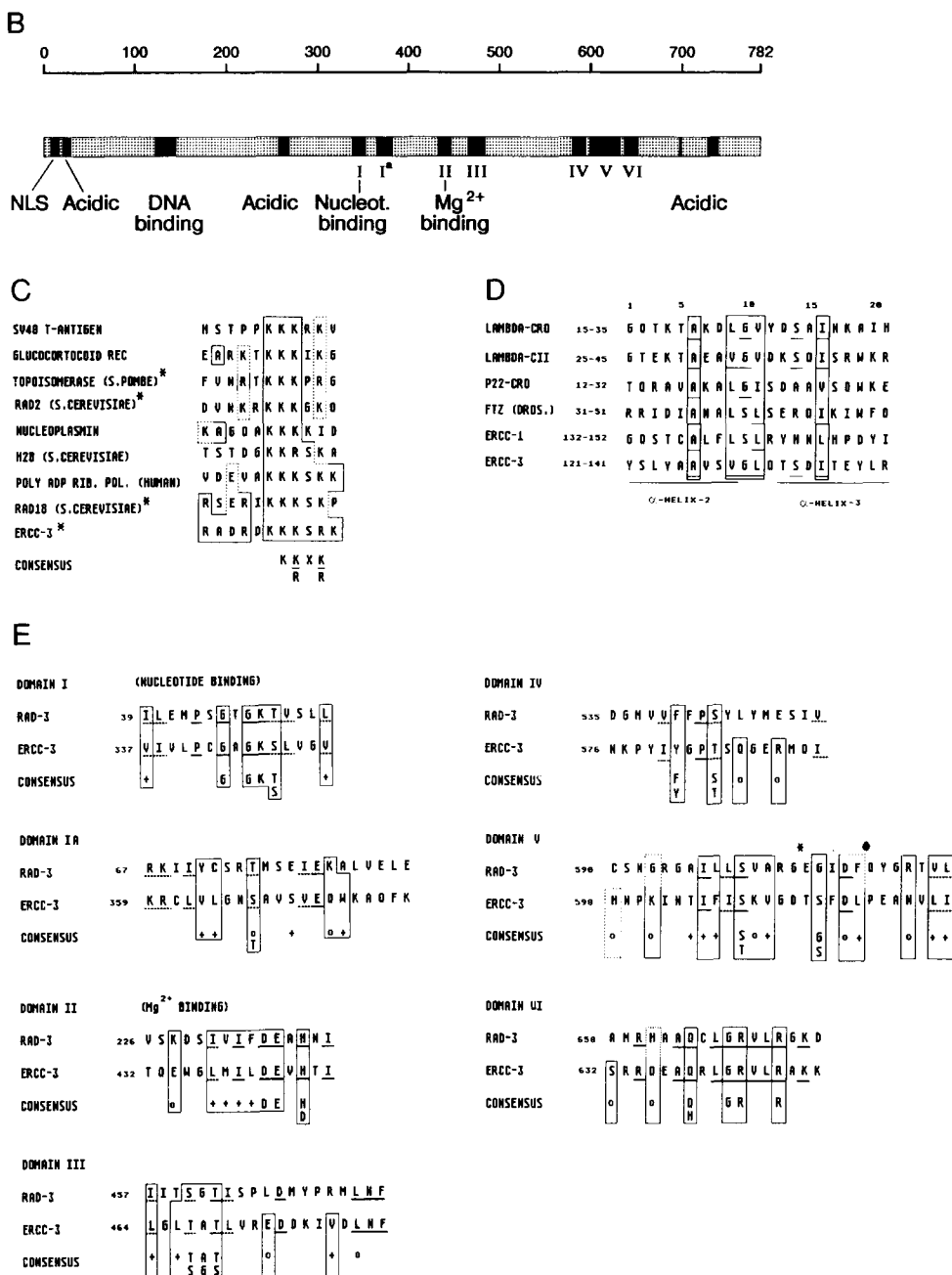


Figure 2. Composite Nucleotide and Encoded Amino Acid Sequences of *ERCC-3* cDNA and Identification of Putative Functional Domains

(A) Composite nucleotide sequence and encoded amino acid sequence of *ERCC-3* cDNA. The 2750 bp sequence was derived from *ERCC-3* cDNA clones pCD1, pHK1, and pHT1. The first ATG, starting a 48 bp ORF, is indicated with "M" and the first stop codon with an asterisk. Amino acid sequence of the long ORF is presented using the one-letter code. Putative functional domains are boxed. Acidic domains are indicated with dashed underlining. The XP11BE mutation is indicated by an arrow. ATTTA regions and the AATAAA polyadenylation signal are underlined. The CAYTG region is double underlined.

(B) Predicted functional domains in the *ERCC-3* protein. NLS, nuclear localization signal.

(C) Amino acid comparison of *ERCC-3* sequences and known or postulated nuclear localization signals of SV40 T antigen, glucocorticoid receptor, *S. pombe* topoisomerase I (Uemura et al., 1987), *S. cerevisiae* RAD2 (Madura and Prakash, 1986) and RAD18 (Jones et al., 1988), poly(ADP-ribose) polymerase (Cherney et al., 1987), nucleoplasm, and *S. cerevisiae* histone H2B. A consensus sequence for nuclear localization signals has been described by Roberts (1989), where references for other sequences can be found. Identical (solid line) or similar (dashed line) residues shared by all proteins are boxed. Nuclear localization signal functions that have not been demonstrated are indicated with an asterisk.

(D) Amino acid comparison of *ERCC-3* and *ERCC-1* and DNA binding domains in *Drosophila fushi tarazu* (FTZ) homeobox proteins (Shepherd et al., 1984) and in prokaryotic activator and repressor proteins λ cro, λ cII, and P22 cro (reviewed in Pabo and Sauer, 1984). See van Duin et al. (1986) for the *ERCC-1* sequence. Amino acids 6, 9-11, and 16, which are important for proper positioning of the two α helices with respect to each other, are boxed. Positions of the two α helices of the DNA binding domain are shown at the bottom. Amino acids are numbered, and those identical with the *ERCC-3* sequences are underlined.

(E) Homology between the amino acid sequences of *ERCC-3* and RAD3 and seven conserved motifs of two superfamilies of more than 20 known DNA and RNA helicases (Gorbalenya et al., 1988, 1989). Identical (solid underline) or strongly homologous (dashed underline) residues shared by the RAD3 and *ERCC-3* proteins are indicated. Charged/polar amino acids (S, T, D, E, N, Q, K, R) are indicated with 0, hydrophobic amino acids (I, L, V, M, F, Y, W) with +. Asterisk indicates insertion of KVS into the RAD3 sequence; closed circle indicates insertion of DH. Amino acids matching the consensus are boxed.

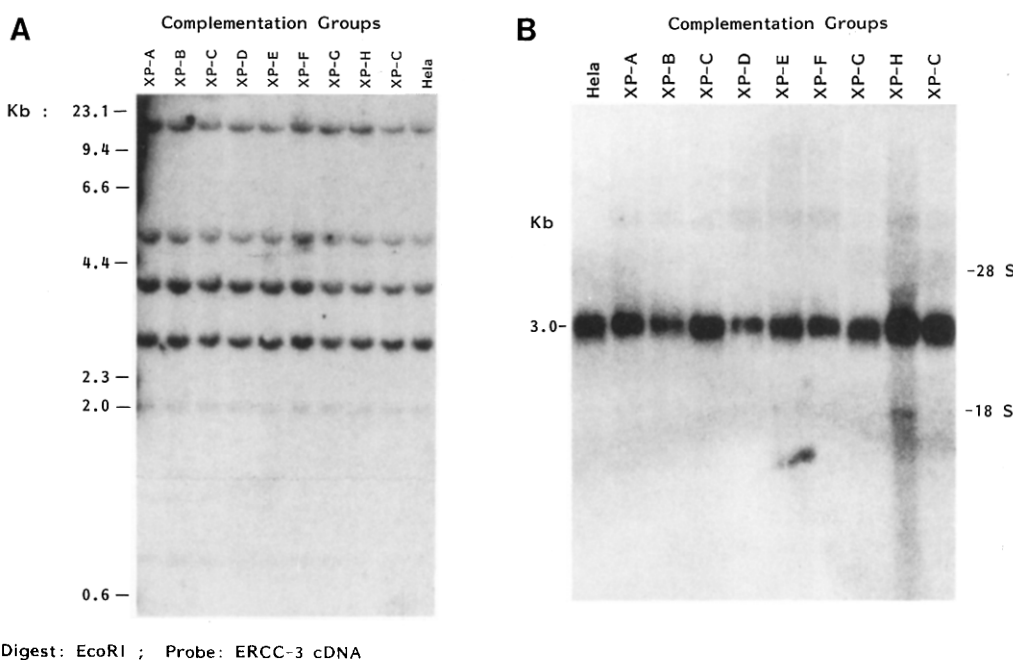


Figure 3. Southern and Northern Blot Analysis of *ERCC-3* in Representative Cell Lines of Eight XP Complementation Groups

(A) Southern blot analysis of 15 µg of EcoRI-digested chromosomal DNA of representative cell lines of all eight XP complementation groups (XP-A to XP-H) and HeLa DNA as a control. The blot was hybridized with a ³²P-labeled *ERCC-3* cDNA probe (insert pCD1), which recognizes all *ERCC-3* exons (unpublished data).

(B) Northern blot analysis of RNA of representative cell lines of all eight XP complementation groups and HeLa RNA as control. Fifty micrograms of total RNA was loaded per lane. The blot was hybridized with an *ERCC-3* probe as described above.

cells. Other enzymes yielded similar results, confirming the absence of major structural abnormalities (data not shown).

Second, to investigate the expression of the endogenous *ERCC-3* gene in a number of XP cell types, total RNA was isolated and analyzed by Northern blot experiments using *ERCC-3* cDNA as a probe (Figure 3B). All eight XP groups examined produced comparable amounts of an *ERCC-3* transcript of the correct length. From these data we conclude that the *ERCC-3* gene in the XP cells tested contains neither mutations affecting the size of the mRNA nor drastically reduced levels of the transcript.

These data do not exclude the possibility that the *ERCC-3* gene is inactivated by a small (point) mutation that impairs its function. To investigate this possibility, functional complementation studies were performed. The *ERCC-3* cDNA expression clone pCD1, driven by the SV40 early promoter (Okayama and Berg, 1982), was cotransfected with the dominant selectable marker pRSVneo into SV40-immortalized XP fibroblasts that are available for six of the XP complementation groups: XP-A, XP-C, XP-D, XP-E, XP-F, and XP-G (see Experimental Procedures). The G418-resistant transformants of each line were grown in mass culture consisting of 20 to more than 100 independent clones. Given the relatively poor transfection properties of many SV40-immortalized human fibroblast lines (Hoeijmakers et al., 1987), we confirmed by Southern and Northern blot analysis that the transfected cDNA was incorporated and

properly expressed by a considerable fraction of the transformants in each of the XP lines (results not shown; see equivalent results for *ERCC-1* in van Duin et al., 1989). To test whether *ERCC-3* restores the DNA repair characteristics, we determined the UV survival of G418-resistant mass populations. All six transformed XP cultures displayed mutant UV survival phenotypes, indicating inability of *ERCC-3* to correct their defects (not shown).

The remaining two XP complementation groups (XP-B and XP-H, both associated with CS), for which no immortalized cell lines suitable for transfection experiments are available, were examined for *ERCC-3* correction by means of microinjection. The *ERCC-3* cDNA construct used above was injected into one of the nuclei of XP homopolykaryons generated by cell fusion of the XP fibroblasts (see Experimental Procedures). To assay for the ability of the injected *ERCC-3* cDNA to correct the excision defect in these XP complementation groups, injected cells were probed for their capacity to perform UV-induced unscheduled DNA synthesis (UDS). To allow expression of the injected gene copies, 18–24 hr after injection the fibroblasts were exposed to UV light (16 J/m²), incubated in the presence of [³H]TdR (2 hr), and processed for autoradiography to visualize incorporation of [³H]TdR in repair patches. Noninjected cells were subjected to the same procedure. The untreated XP-B (XP11BE) fibroblasts characteristically showed a strongly reduced level of UDS, indicated by a low number of autoradiography grains above their

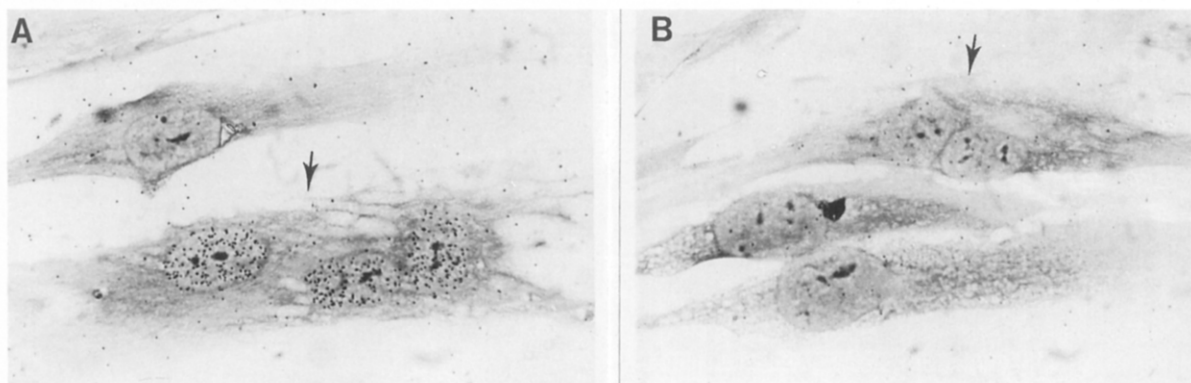


Figure 4. Microinjection of *ERCC-3* cDNA Corrects UV-Induced UDS in XP-B Fibroblasts

Micrographs of XP11BE (complementation group B) homopolykaryons microinjected with the *ERCC-3* expression plasmids in one of the nuclei, followed by incubation for 18 hr to permit expression of the injected cDNA, UV irradiation, and assay for UDS. (A) Homopolykaryon injected with pSVH3. (B) Homopolykaryon injected with pSVH3M. Injected cells are indicated with arrows.

nuclei (Figure 4, noninjected cells; see also Table 1). The XP-H fibroblasts displayed a considerable residual UDS. As shown in Figure 4A and Table 1 (experiments I and IV), injection of *ERCC-3* cDNA did clearly restore UDS in the XP11BE homopolykaryons to wild-type levels and had no effect on UDS of the XP-H cells.

Several control experiments were carried out to validate this finding. The specificity of the correction was established by injection of *ERCC-3* cDNA into fibroblasts of other XP complementation groups. No correction was found in XP-G or XP-D cells, in agreement with the transfection results (data not shown). On the other hand, injection of *ERCC-1* cDNA or of an incomplete and inactive *ERCC-3* cDNA construct (pCD2) into XP11BE did not significantly

enhance the UDS in XP11BE (Table 1; van Duin et al., 1989). Furthermore, when UV irradiation was omitted, no increase in the number of grains was observed, ruling out the possibility that the grains were due to damage-independent DNA synthesis. The XP11BE correction is very powerful. Dilution experiments demonstrated that injection of a single *ERCC-3* cDNA molecule into one of the nuclei of a homopolykaryon still induced wild-type UDS in all nuclei of the injected cell (not shown). Finally, cytoplasmic injection of in vitro synthesized *ERCC-3* RNA transcribed from the pEP-1 construct (a *ERCC-3* cDNA plasmid containing the SP6 promoter; see Experimental Procedures) also enhanced UV-induced UDS of XP-B fibroblasts (Table 1, experiment III). We conclude that func-

Table 1. Effect of Injection of *ERCC-3* cDNA Expression Plasmids and In Vitro Synthesized *ERCC-3* RNA on the UV-Induced UDS of XP Fibroblasts

Experiment	Injected Cell ^a	Sample	Expression Time (hr)	UV Irradiation	UDS (grains/nucleus)	% of UDS Relative to Wild Type
I ^b	XP11BE (B)	pCD1	24	(+)	257 ± 3	104
		pCD2	24	(+)	12 ± 3	5
		pCD1	24	(-)	2 ± 1	—
		—	—	(+)	11 ± 3	5
II ^b	XP11BE (B)	pSVH3	24	(+)	124 ± 10	94
		pSVH3M ^c	24	(+)	6 ± 1	5
		—	—	(+)	6 ± 1	5
III ^d	XP11BE (B)	RNA	6–8	(+)	81 ± 8	68
		RNA	18–20	(+)	85 ± 7	75
		—	—	(+)	6 ± 2	5
IV ^b	XP2CS (H)	pCD1	24	(+)	50 ± 2	41
		pCD1	24	(-)	1 ± 0	—
		—	24	(+)	52 ± 3	43

In each experiment more than 20 nuclei of homopolykaryons were injected and counted. The average number of grains per (non-S-phase) nucleus is given. In every experiment the repair-proficient fibroblast cell line C5RO was included as an internal standard and positive control.

^a XP complementation group indicated in parentheses.

^b DNA concentration: 0.5 µg/µl.

^c A 100-fold excess did not change the UDS levels.

^d RNA concentration: 0.2 µg/µl.

tional *ERCC-3* cDNA very specifically and efficiently corrects the excision repair defect in XP11BE, strongly suggesting that the *ERCC-3* gene is the defective gene in this XP/CS complementation group.

Determination of the *ERCC-3* Mutation in XP11BE Cells

Since the *ERCC-3* gene is not detectably rearranged in XP11BE and its RNA is properly expressed (see Figure 3), it is most likely inactivated by a small-scale mutation. To search for such a mutation, we utilized the polymerase chain reaction (PCR; Saiki et al., 1988) to amplify cDNA from mRNA of cells (fibroblasts and lymphoblastoid lines) from the sole XP-B patient (XP11BE) and her mother. Since the *ERCC-3* transcript is too large and too rare to be amplified in one step, its mRNA was amplified in three segments: 5' (706 bp), central (899 bp), and 3' (1219 bp). Sequence analysis revealed that the 3' section of the patient's cDNA had incurred a 4 bp (GCAG) insertion at position 2220, disturbing the ORF (Figure 5). No changes were observed in the remainder of the XP11BE cell *ERCC-3*. The deviation from the normal sequence was found in all eight independent cDNA clones of XP11BE sequenced.

To determine whether this mutation is inherited, the relevant part of the cDNA of the patient and her mother (no cells of the father are available) were amplified and subjected together with cDNA from HeLa cells to dot blot analysis. The filters were hybridized with ³²P-labeled oligonucleotides specific for the wild-type and mutant genes (Figure 6). As an internal control on the quality of the RNA, cDNA synthesis, and amplification, amplimers specific for the *N-ras* oncogene were included in the reaction. The specificity of the hybridization was tested using the cloned wild-type and mutated *ERCC-3* cDNA spotted on the same filter. Hybridization of the *N-ras* probe demonstrates that all RNAs gave amplification of the *N-ras* transcripts (Figure 6C). As expected, the HeLa control displays only hybridization to the wild-type-specific *ERCC-3* oligonucleotide probe (Figures 6A and 6B). RNA of the patient shows clear hybridization with the mutant-specific probe (Figure 6B) and hardly any signal with the wild-type-specific probe (note that this probe gives a weak cross-hybridization to the mutant *ERCC-3* cDNA; Figure 6E, right). The RNA from the mother, however, exhibits clear hybridization with both probes (Figures 6A and 6B). This indicates that she is heterozygous for the mutation in the *ERCC-3* gene and that the patient inherited this insertion from her mother. The sequence analysis of cDNA clones, the dot blot hybridization, and S1 nuclease analysis (not shown) indicate that most, if not all, *ERCC-3* transcripts in the XP11BE cells carry the 4 bp insertion.

Mutation at the Genome Level Is a C→A Transversion in a Splice Acceptor Site

What is the molecular mechanism responsible for the insertion of the GCAG sequence at the mRNA level? The last 3 bases of the insertion correspond to the consensus sequence for splice acceptor sites (CAG) found at the end of introns. This suggested that improper splicing might account for the alteration at the level of mature *ERCC-3*

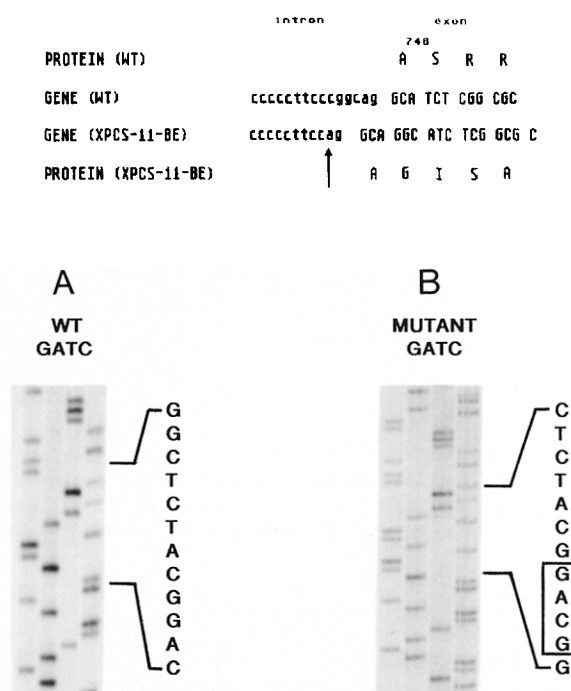


Figure 5. Nucleotide Sequences of the Patient XP11BE (Mutant) and HeLa (Wild-Type) Alleles of *ERCC-3* cDNA and Gene

(Top) Nucleotide sequence of the 3' end of the intron (lowercase) and the adjacent exon (capitals). The arrow indicates the C→A transversion in the genomic DNA of the XP11BE patient (maternal allele). The wild-type and mutant predicted amino acid sequences are given in the one-letter code.

(Bottom) Part of a wild-type *ERCC-3* cDNA sequencing gel (A) and the corresponding region in the XP11BE cDNA (B) at the mutated site. The 4 bp insertion is boxed.

mRNA. To resolve the genetic basis of the mutation and to see whether the patient is heterozygous for it, we examined the *ERCC-3* gene in XP11BE and HeLa chromosomal DNA. The genomic organization of *ERCC-3* has been elucidated for the most part. Unfortunately, the portion of the gene in which the mutation resides has so far been refractory to cloning in *E. coli*: we have not been able to subclone it from cosmid and λ libraries covering ≥ 60 human genomic equivalents (Weeda et al., 1990).

We therefore decided to amplify the genomic region surrounding the mutation by PCR with the help of the flanking primers p34 and p39 (see Figure 1). This yielded several clones from both the XP11BE and HeLa genomic DNA. Sequence analysis of the HeLa (wild-type) *ERCC-3* gene between the primers revealed the presence of a splice acceptor site exactly at the position of the mutation in XP11BE transcripts (Figure 5). In the XP11BE DNA two sequences were found: one identical with the wild-type sequence, the other containing a C→A transversion at a position that creates a new splice acceptor site consensus sequence 4 bp upstream of the one used in the wild-type situation. This is expected to shift the intron/exon border 4 bp into the intron, resulting in the incorporation of an additional 4 bp into the adjacent exon. The first allele must represent the paternal gene copy, the second that origi-

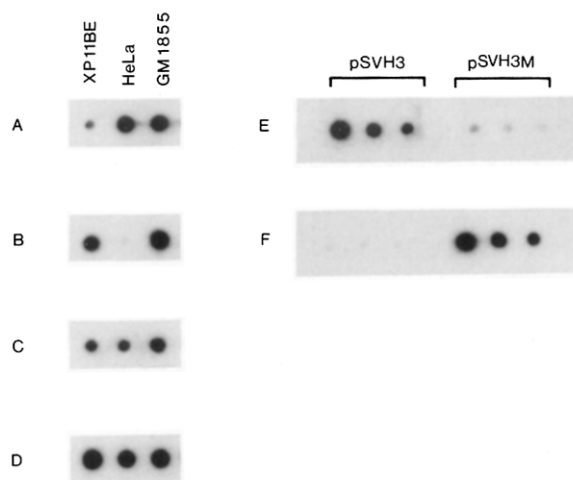


Figure 6. PCR Amplification of *ERCC-3* cDNAs between Primers p34 and p35

See Figure 1B for amplification strategy. Amplified DNA from HeLa cells, XP11BE, and GM1855 (mother of XP11BE) was spotted onto Bio-trace nylon filters and hybridized with a labeled wild-type oligonucleotide probe (p42; A and E) or a mutant oligonucleotide probe (p43; B and F) (see Experimental Procedures for primer sequences). To ensure that the samples are equally capable of undergoing amplification of specific sequences by the PCR method, amplimers specific for the *N-ras* gene (around codon 12) were also added to the cDNA mix and hybridized to an *N-ras* 12-specific oligonucleotide probe (C). To confirm that equal amounts of amplified material were spotted onto the filter, the filter was hybridized with an *ERCC-3*-specific oligonucleotide (p12; D). As a control for the specificity of hybridization of mutant and wild-type *ERCC-3*-specific oligonucleotides, plasmid DNA from clones pSVH3 (wild type) and pSVH3M (XP-B mutation) was spotted in serial dilution (E and F). Note that the wild-type-specific *ERCC-3* oligonucleotide probe (p42) weakly cross-hybridizes to the mutant cDNA in pSVH3M (E, right).

nating from the mother. The sequence predicts that in the mutant *ERCC-3* allele there should be a BstNI restriction site at this position and none in the paternal allele.

To confirm the results described above, we subjected chromosomal DNA of XP11BE and her mother to BstNI restriction digestion (Figure 7) and performed Southern blot analysis using a ³²P-labeled probe that recognizes only the last exon. As anticipated, BstNI degraded approximately 50% of the wild-type fragment from the patient and her mother. Collectively, these results imply that the patient and her mother are both heterozygous for a splice mutation inactivating one *ERCC-3* allele, and that the paternal allele in the patient must contain another small mutation causing the *ERCC-3* mRNA derived from this gene copy to be absent or greatly reduced.

XP Mutation Inactivates the *ERCC-3*

Repair Function

The extra 4 bp present in the *ERCC-3* mRNA of XP11BE causes a frameshift starting at codon 740. To determine the effect of this C-terminal mutation on *ERCC-3* function, a cDNA construct (pSVH3M) carrying the 4 bp insertion was made and was microinjected into XP11BE fibroblasts. No significant enhancement of UDS was found even at

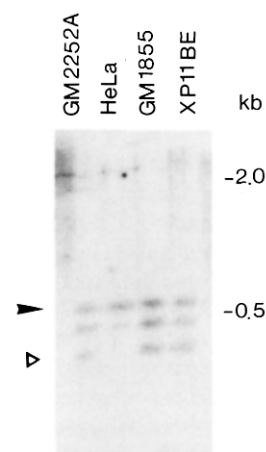


Figure 7. Demonstration of a Point Mutation in One of the *ERCC-3* Alleles of XP11BE by Southern Blot Analysis

Genomic DNA from XP11BE (fibroblast cell line of the patient), GM2252A (lymphoblastoid cell line of the patient), GM1855 (lymphoblastoid cell line of the mother), and HeLa cells was digested with BstNI and electrophoresed on a 1.2% agarose gel and blotted onto a nylon filter. The filter was hybridized with a ~400 bp *ERCC-3* cDNA Nari-coRI fragment (see Figure 1; EcoRI restriction site is part of the linker sequence) containing the last exon. Owing to cross-hybridization to thus far unknown sequences (possibly derived from a pseudogene), two fragments hybridize in the HeLa control. Sequence data of the last intron and 3' flanking sequences of the *ERCC-3* gene (unpublished) indicate that the upper fragment (indicated with an arrowhead) corresponds to the last *ERCC-3* exon. The middle fragment is derived from the cross-hybridizing sequence. The open triangle denotes the fragment specific for the XP-B mutation.

cDNA concentrations 100-fold higher than required for the normal cDNA to yield wild-type correction (Figure 4B and Table 1). Furthermore, transfection of pSVH3M into the CHO mutant 27-1 failed to increase the UV survival of this rodent repair mutant (not shown). These experiments indicate that the XP11BE mutation inactivates the *ERCC-3* repair function, as evidenced by its inability to correct the UV survival of 27-1 cells and the UDS of XP11BE itself.

Discussion

This paper identifies the molecular defect in a human DNA repair disorder. A C→A transversion in the excision repair gene *ERCC-3* was shown to underlie the inborn defect in XP complementation group B, a complex and remarkable conjunction of xeroderma pigmentosum and Cockayne's syndrome. The gene that we hence designate XP-B-correcting (*XPBC*)/*ERCC-3* is predicted to encode a DNA repair helicase. In addition, this study reveals an overlap between the class of rodent mutants and naturally occurring human inborn defects of DNA repair. This finding demonstrates the importance of the rodent repair mutants as a model for human repair disorders.

Combination of XP and CS

XP11BE, the sole patient representing XP-B, was the only affected child in a family consisting of seven daughters of

unrelated parents. She remains remarkable as she clearly manifests the clinical symptoms of two rare excision repair disorders: CS (sensitivity to sunlight, dwarfism, microcephaly, wizened facial appearance, deafness, severe mental retardation) and XP (extreme sensitivity to sunlight, pigmentation abnormalities, predisposition to skin cancer [the latter two are not normally associated with CS]) (Cleaver and Kraemer, 1989). The patient's cellular defects include hypersensitivity to UV light and virtually complete absence of excision repair as measured by the level of UV-induced UDS (which also is unusual for CS) (Lehmann, 1987). At the biochemical level, CS cells are thought to be disturbed in the preferential repair of actively transcribed genes, their overall genome repair being normal (Venema et al., 1990b). XP patients harbor deficiencies affecting both processes or, in the case of XP-C, only the "global" genome repair subpathway (Venema et al., 1990a).

In view of the extreme rarity of both disorders (of the order of 10^{-5} each), it is highly unlikely on statistical grounds that XP11BE represents a double mutation affecting two distinct genes. In this patient's family no siblings suffering from either of the syndromes have occurred, nor has any patient been identified that falls into the same complementation group and shows the clinical features of only one of the diseases. However, the possibility that two closely linked loci are involved, both of which are inactivated by one deletion, cannot be dismissed on the basis of these arguments. The Southern and Northern blots and sequence analysis data, as well as the microinjection results presented here, indicate that the *XPBC/ERCC-3* gene fully and specifically corrects the repair defect in XP-B and that it is inactivated by small-scale mutations. This rules out the involvement of additional genes by way of gross gene rearrangements. Hence, the *XPBC/ERCC-3* protein is either implicated in two excision repair subpathways, one associated with XP and the other with CS, or—more likely—participates in a common step leading to the phenotypic features of both disorders (Lehmann, 1987). Since inactivation of *XPBC/ERCC-3* results in predisposition to skin cancer, the gene can functionally be considered a tumor-preventing gene.

The Mutation in the *XPBC/ERCC-3* Gene

Analysis of the *XPBC/ERCC-3* gene in the XP11BE patient's chromosomal DNA revealed that her maternal allele harbored a C→A transversion in the last intron, generating an optimal splice acceptor sequence 4 bp upstream of the normal 3' splice site. In vitro experiments have indicated that the spliceosome complex scans for the first AG dinucleotide located 3' of the branchpoint and a polypyrimidine tract (Smith et al., 1989). In some instances where a series of closely spaced AG sequences occur, subsequent AGs may occasionally also be interpreted as the start of the next exon. Nevertheless, in these instances the first AG is the one predominantly used (Fu et al., 1988). Sequence analysis of PCR-generated *XPBC/ERCC-3* cDNA clones of XP11BE RNA and dot blot data demonstrate that the vast majority, if not all, of the *XPBC/*

ERCC-3 transcripts contain the intron-derived 4 bp insert and, therefore, used the first (C)AG.

These results provide an in vivo confirmation of the above findings with the in vitro splicing studies based on nuclear extracts. Our data agree with those for the behavior of a number of splicing mutants in other inherited disorders in which new coding segments are inserted or in which protein function is disrupted by in-frame insertion (Nakano and Suzuki, 1989) or by frameshifting of the normal exon sequences (Spritz et al., 1981; Metherall et al., 1986). The parents of XP11BE are unrelated; hence it is not unexpected that the patient appeared to be heterozygous for the splice mutation. Since we failed unequivocally to detect transcripts without the 4 bp insertion, we must presume that the paternal allele is not detectably expressed at the level of mature *XPBC/ERCC-3* mRNA. This may be due to a mutation affecting either the promoter of the gene or the processing, transport, or stability of its transcript.

The DNA Repair Defect in XP11BE

Although XP11BE fibroblasts have a greatly reduced level of repair synthesis, they still display detectable UDS (see Table 1). This low residual activity may be explained in various ways. First, we cannot completely rule out the possibility that a small amount—too low to be detected by our techniques—of functional *XPBC/ERCC-3* transcript is synthesized. Such RNA could be derived from the paternal allele, or, if its origin is maternal, it could still be a rare, correctly spliced molecule. Second, it is possible that the mutated *XPBC/ERCC-3* gene product still fulfills part of its functions or displays a dose-independent, severely decreased activity. In this respect it is worth noting that the frameshift affects only the C-terminal 41 amino acids of the predicted *XPBC/ERCC-3* protein, a region as yet devoid of assigned functional domains. Finally, it can be envisaged that the residual UDS originates from other repair processes or from excision repair subpathways that are not completely blocked by the lack of functional *XPBC/ERCC-3* protein. Further investigations are required to discriminate between these alternatives.

The Presumed Helicase Function of *XPBC/ERCC-3*

Systematic comparison of the predicted amino acid sequence of the *XPBC/ERCC-3* gene product with known functional domains in other proteins suggests that it may be a DNA, chromatin, and nucleotide binding helicase. The deduced protein bears a striking resemblance (almost perfect match and proper spacing) to seven consecutive amino acid motifs identified in two related superfamilies of putative or already established helicases. The set of helicase domains invariably starts with a nucleotide binding region. In addition, we have identified a stretch with structural homology to "helix-turn-helix" DNA binding motifs and to various very acidic domains encountered, e.g., in the high mobility group I protein family, which is thought to carry the ability to associate with basic histones in chromatin (Reeck et al., 1982).

Since DNA unwinding constitutes a key step in the initi-

ation of many processes of nucleic acid metabolism, it is not surprising that helicases are increasingly recognized to be important for many fundamental cellular systems. Thus members of this rapidly expanding family of proteins are implicated in the initiation and progression of DNA replication (Brill and Stillman, 1989) and transcription (Sopata et al., 1989), as well as cell proliferation (Ford et al., 1988; Hay et al., 1988), splicing (S  raphin et al., 1989), ribosome assembly (Nishi et al., 1988), and initiation of translation (Abramson et al., 1987). Furthermore, these enzymes play a central role in DNA recombination (e.g., Kodahek and Alberts, 1987) and repair. In *E. coli* the *uvrAB* complex has been shown to exhibit DNA helicase activity, which is used for scanning of the DNA for structural perturbations and/or local unwinding of the DNA strands at the site of the lesion (Oh and Grossman, 1987). At later stages, when the *uvrC* endonuclease has introduced two scissions into the damaged strand, the *uvrD* protein (i.e., helicase II, a 3'→5' unwinding protein) is involved in the release of the lesion-containing oligonucleotide from the DNA (Matson, 1986; see Sancar and Sancar, 1988, for a review).

Sung and coworkers (1987a, 1987b) have demonstrated that the yeast excision repair protein *RAD3* also encodes an ATP-dependent (5'→3') DNA helicase. Interestingly, Weber et al. (1990) have found recently that the predicted human excision repair protein *ERCC-2* harbors a ≥50% overall sequence similarity to the *RAD3* polypeptide. This implies that *ERCC-2* is the human equivalent of yeast *RAD3*. All essential amino acids in the seven helicase domains are strongly conserved between the yeast and human proteins, suggesting that the human gene product may also be a 5'→3' DNA helicase.

Our findings identify a second putative DNA unwinding protein in mammalian excision repair. It is tempting to speculate that the *ERCC-2* and *XPBC/ERCC-3* proteins may fulfill closely associated, equivalent, or complementary functions. This is consistent with the fact that the phenotypes of mutant rodent groups 2 and 3 are very similar: both have a similar degree of sensitivity to UV and agents causing bulky DNA adducts, no pronounced cross-sensitivity toward X-rays and DNA cross-linking agents such as mitomycin C, a similar level of UV-induced mutagenesis, and a defect in early (pre)incision steps of the excision repair pathway (Thompson et al., 1980, 1982; Thompson, 1989; Zdzienicka and Simons, 1987; Zdzienicka et al., 1987). Furthermore, we have identified a yeast homolog of *XPBC/ERCC-3* that exhibits a level of amino acid similarity to *XPBC* as high as that found between the *ERCC-2* and *RAD3* proteins (Koken, Weeda, and Hoeijmakers, unpublished data).

These considerations are consistent with the hypothesis that *ERCC-2* and *XPBC/ERCC-3* exert a similar function. An obvious possibility is that the *XPBC/ERCC-3* protein unwinds DNA in the direction opposite (3'→5') to that of the *RAD3*-type helicase (e.g., *ERCC-2*) and that both work together in locally unwinding the duplex DNA on both sides of the lesion, as indicated in the model of Figure 8. Alternatively, it is possible that the two enzymes cooperate in scanning long stretches of DNA for structural distortions or in ridding the DNA around the lesion of nucleosomes, to give repair complexes access to the site of the DNA injury.

One final point concerns another important function of the *RAD3* gene, and presumably the *ERCC-2* gene as well. Null alleles of *RAD3* generated by gene disruption or point mutations appear to be recessively lethal, indicating an essential role for *RAD3* in yeast cell viability (Higgins et al., 1983; Naumovski and Friedberg, 1983). The notably low recovery of group 2 (*ERCC-2*) rodent mutants as the result of mutagenesis with mutagens that cause mainly frameshifts as compared with agents that induce predominantly point mutations strongly suggests that frameshifts in the *ERCC-2* gene tend to be lethal in the hamster as well (Busch et al., 1989). It could be that the *XPBC/ERCC-3* gene, by analogy with *ERCC-2* and *RAD3*, has a vital function in addition to its role in DNA excision repair. The proposition that only those mutations in the *XPBC/ERCC-3* gene are tolerated that leave its vital function intact could explain the extreme rarity of this type of inborn defect and our finding that the mutation in the expressed allele is situated in the very C-terminal part, outside of the regions to which we have assigned specific functions. The issue of whether *XPBC/ERCC-3* has an essential role in cell viability can be more easily addressed in the yeast system, and these experiments are currently under way.

Experimental Procedures

General Procedures

Purification of nucleic acids, restriction enzyme digestion, gel electrophoresis, nick translation, and filter hybridization were performed according to established procedures (Maniatis et al., 1982).

Cell Lines and Transfection Experiments

UV-sensitive 27-1 cells (Wood and Burki, 1982) were cultured in F10/Dulbecco's minimal essential medium (DMEM) supplemented with 5% newborn calf serum and 5% fetal calf serum (GIBCO), 100 µg/ml streptomycin, and 100 U/ml penicillin. XP fibroblasts were cultured in F10 supplemented with 10% fetal calf serum. Transformed cell lines harboring the *E. coli gpt* dominant marker gene were selected and cultured in XGPT medium containing F10/DMEM as described above supplemented with 0.2 µg/ml aminopterin, 5 µg/ml thymidine, 10 µg/ml xanthine, 25 µg/ml mycophenolic acid (GIBCO), 15 µg/ml hypoxanthine, and 2.3 µg/ml deoxycytidine. Cell lines containing the selectable

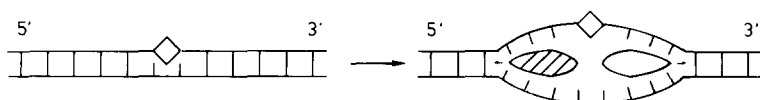


Figure 8. Model for the Action of Repair Helicases

The repair helicases *ERCC-2* and *XPBC/ERCC-3* (hatched) are proposed to work together to unwind the duplex DNA at both sites of the lesion in an early preincision step of the excision repair pathway.

marker aminoglycosyl phosphotransferase were selected with the antibiotic G418 (100–400 µg/ml, depending on the cell line). The selection medium was refreshed every 3–5 days.

Transfection of 27-1 and SV40-immortalized XP fibroblasts was performed essentially as described by Graham and van der Eb (1973). XP cells were transfected at 80% confluence with 1 µg of pRSVneo plus 2 µg of the *ERCC-3* expression plasmid pCD1; 27-1 cells were transfected with 1 µg of pSV3gptH (Westerveld et al., 1984) plus 2 µg of pCD1. After transfection the cells were shocked in 5 ml of PBS containing 10% dimethyl sulfoxide for 1–2 min, and rinsed twice with 10 ml of PBS. The medium was refreshed 48 hr after transfection by medium containing 100–400 µg/ml G418 (GIBCO) or XGPT medium. Selection for repair-proficient transformants was done as follows: after the appearance of mycophenolic acid- or G418-resistant colonies, each dish was trypsinized and the cells were reseeded into two dishes and irradiated with 254 nm (peak) germicidal UV light at a flux of 0.5 J/m². Cells were irradiated three times with 4.5 J/m² UV light at 1 day intervals. For the primary and SV40-transformed XP fibroblasts lines used in this study, see van Duin et al. (1989) and references therein. GM1855 (lymphoblastoid cell line of XP11BE) and GM2252A (lymphoblastoid cell line of mother of XP11BE) were generously provided by Dr. K. Kraemer (National Institutes of Health, Bethesda, MD).

Survival Experiments

For UV survival experiments, exponentially growing cultures were trypsinized and 10²–10⁴ cells were plated onto 6 cm dishes and left to attach for about 12 hr. Subsequently, cells were rinsed with PBS and exposed to UV light (see above). After cultivation in nonselective medium for 7 days, clones were fixed and stained with Coomassie brilliant blue.

Microinjection into Fibroblasts

XPB/ERCC-3 cDNA constructs pCD1 and pCD2, with or without addition of pSV3gpt in a 1:1 molar ratio (0.5 µg/ml, final concentration), were injected into the nuclei of XP homopolykaryons as described by van Duin et al. (1989) using procedures previously described (de Jonge et al., 1983). Cells were incubated for 24–48 hr to allow expression of the injected DNA. To assess the success rate of injection, one part of the injected cells (in the coinjection experiments) was stained for the expression of SV40 large T antigen encoded by pSV3gpt, using a monoclonal antibody and a fluorescein isothiocyanate-containing goat antimouse conjugate. To determine the effect on the repair potential, the other part of the injected cells was subjected to UV-induced UDS, which was measured by autoradiography as described previously (Vermeulen et al., 1986). In vitro synthesized RNA was injected into the cytoplasm, and the cells were assayed for UDS 6–20 hr after injection.

Southern Blot Analysis

Restriction endonuclease digestions were carried out as recommended by the manufacturer (Pharmacia). Chromosomal DNA (15 µg) was size fractionated by gel electrophoresis on a 0.8% agarose gel. After alkaline treatment the DNA was blotted onto nylon filters (Biotrace) according to the supplier's instructions. Probes were labeled by nick translation.

RNA Isolation and Northern Blot Analysis

Total RNA was isolated by the LiCl-urea method described by Auffray and Rougeon (1980). Cytoplasmic RNA was isolated by the NP40 method as described in Maniatis et al. (1982). RNA samples were separated by electrophoresis through a 1% agarose-formaldehyde gel and transferred onto nitrocellulose. The filters were hybridized at 42°C in formamide (50% [v/v]), 5× SSPE, 5× Denhardt's solution, 1% SDS, and 100 mg/ml denatured salmon sperm DNA. The filters were hybridized with a ³²P-labeled nick-translated probe. Filters were washed two times with 2× SSC containing 0.1% SDS at 60°C.

Isolation of cDNA Clones and Sequencing

The 1.5 kb PstI insert fragment of pCD1 was used as a probe for isolating other cDNA clones of human origin. To this end, a λgt10 human testis cDNA library (Clone Tech) and a cDNA library of human chronic myelogenous leukemia cell line K562 were screened. Positive λ clones HT3 and HK1 were subcloned into M13mp18 or M13mp19 vectors, respectively, for sequence analysis by the dideoxy chain termination

method (Sanger et al., 1977) with Sequenase (United States Biochemical Corp.) and M13 or sequence-specific primers, or cloned into appropriate expression vectors to yield pH3T, pH3T1, and pHK1, respectively. Sequences were analyzed and compared using the Wisconsin software and the FASTA program.

DNA Amplification and Mutation Detection

Total or cytoplasmic RNA (10 µg) was used for preparing cDNA with *ERCC-3*-specific primers (see below). RNA was dissolved in 9 µl of annealing buffer (250 mM KCl, 10 mM Tris-HCl [pH 8.3], 1 mM EDTA). Following the addition of 1 µl (100 pmol/µl) of primers, the samples were first heated for 3 min at 80°C and transferred to a 37°C water bath for 1 hr, 15 µl of cDNA buffer (24 mM Tris-HCl [pH 8.3], 16 mM MgCl₂, 8 mM DTT, 0.4 mM of dGTP, dATP, dTTP, and dCTP) and 5 U of Moloney murine leukemia virus reverse transcriptase (Promega) were added, and the tubes were incubated at 37°C for 1 hr. To 5 µl of cDNA, 10 µl of Taq buffer (100 mM Tris-HCl [pH 8.3], 15 mM MgCl₂, 500 mM KCl, 2 mg/ml BSA), 4 µl of dNTPs (2.5 mM), 75 µl of sterile water, 1 µl (100 pmol/µl) of each primer, and 2 U of Taq polymerase (Cetus) were added. Amplification was performed by 32 consecutive cycles of 1 min at 95°C, 1 min annealing at 50°C, and 3 min of extension at 70°C as described by Saiki et al. (1985). Chromosomal DNA was amplified as described above. Amplified DNA was spotted onto Nylon (Biotrace) filters, UV cross-linked, and hybridized to wild-type and mutant ³²P-labeled primers. As an internal control, the N-ras allele around codon position 12 was amplified from the same RNA sample, as described by Verlaan-de Vries et al. (1986). The amplified DNA was purified and digested with the appropriate enzyme or treated with mung bean exonuclease as indicated in Figure 5 and subsequently cloned into M13mp18 or M13mp19 vectors. Oligonucleotide primers for cDNA, DNA amplification, and DNA sequencing were synthesized in an Applied Biosystems DNA synthesizer. The PCR primers used in this study are described below:

p12: 5'-GGGTGTACATGTTTGC-3'
p33: 5'-GGATCCACCATGGGCAAAAGAGACCG-3'
p34: 5'-GCGCAAGCTTACTGGTTCATAGTGCATC-3'
p35: 5'-CTGATTGGGCCTAAGCTCTAC-3'
p36: 5'-CCTCAGCACACTGGACTTTG-3'
p37: 5'-GCGCGAATTCGGATTGAGCCGGAAGTCCCC-3'
p39: 5'-GCGGATCCTGATCAGGAACTCGCTG-3'
p40: 5'-GGGAGGCCACTGAGCTCATC-3'
p41: 5'-CGGGAAGTGGAGGGCCACC-3'
p42: 5'-CCGAGATGCCTGGCTGGA-3'
p43: 5'-GAGATGCCTGCCTGGCTG-3'
p44: 5'-GACATAGAACTCATGGTGCC-3'
p45: 5'-GGTGGTGGCTGGGGAATTG-3'
p55: 5'-GACTAACATGGGCTGGTTCC-3'

In Vitro Transcription

RNA was synthesized in a 100 µl reaction in which 2 µg of linearized substrate DNA was incubated for 1 hr at 37°C with SP6 RNA polymerase in the presence of RNasin (Promega). The RNA was purified and ethanol precipitated. The RNA was capped as recommended by Promega. After purification, it was microinjected into the cytoplasm of homopolykaryons of XP-B fibroblasts and the effect on induced UDS determined, as described above.

Plasmids

The plasmid used in the study for synthesizing RNA was constructed as follows: pCD1 was used to mutate the -3 position (G→A) in front of the translation initiation codon by PCR, using an oligonucleotide primer containing a BamHI restriction site (using the oligonucleotide primer p33). The 2.7 kb BamHI cDNA insert of pCD1 (containing a poly[A] tract) was cloned into pSP65, yielding pEP1. The cDNA expression plasmids for microinjection were constructed as follows: pSVH3 contains the 5' 1.5 kb EcoRI-PstI fragment of the ORF of pH3T1 ligated together with the 3' 1.2 kb PstI-HindIII fragment of pHK1 into a modified pSVL eukaryotic expression vector (Pharmacia). In pSVH3M, the 3' part was replaced by a 1.2 kb PstI-HindIII fragment synthesized by means of amplifying XP-B cDNA (with oligonucleotide primers p34 and p35).

Acknowledgments

We are grateful to Dr. T. Berkvens, M. Koken, and Dr. K. Kraemer for helpful discussion, Dr. H. van Ormondt for critically reading this manuscript, B. Klein (Leiden) for providing several immortalized XP cell lines, and Dr. G. Grosveld (Rotterdam) for donating the K562 cDNA library. We are indebted to Dr. H. Vrieling (Leiden) for helpful discussions and for help in the initial PCR experiments. M. Kuit is acknowledged for photography. This work was supported by the Netherlands Organization for Advancement of Pure Research (NWO) through the Foundation of Medical Scientific Research (contract no. 900-501-091) and by EURATOM (contract no. BJ6-141-NL).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received May 30, 1990; revised June 29, 1990.

References

- Abramson, R. D., Dever, T. E., Lawson, T. G., Ray, B. K., Tach, R. E., and Merrick, W. C. (1987). The ATP-dependent interaction of eukaryotic initiation factors with mRNA. *J. Biol. Chem.* **262**, 3826–3832.
- Arrand, J. E., Bone, N. M., and Johnson, R. T. (1989). Molecular cloning and characterization of a mammalian excision repair gene that partially restores UV resistance to xeroderma pigmentosum complementation group D cells. *Proc. Natl. Acad. Sci. USA* **86**, 6997–7001.
- Auffray, C., and Rougeon, F. (1980). Purification of mouse immunoglobulin heavy-chain messenger RNAs from total myeloma tumor DNA. *Eur. J. Biochem.* **107**, 303–314.
- Berget, S. M. (1984). Are U4 small nuclear ribonucleoproteins involved in polyadenylation? *Nature* **309**, 179–181.
- Brill, S. J., and Stillman, B. (1989). Yeast replication factor-A functions in the unwinding of the SV40 origin of DNA replication. *Nature* **342**, 92–95.
- Busch, D., Greiner, C., Lewis, K., Ford, R., Adair, G., and Thompson, L. H. (1989). Summary of complementation groups of UV-sensitive CHO cell mutants isolated by large-scale screening. *Mutagenesis* **4**, 349–354.
- Cherney, B. W., Wesley McBride, O., Chen, D., Alkhatib, H., Bhatia, K., Hensley, P., and Smulson, M. E. (1987). cDNA sequence, protein structure, and chromosomal location of the human gene for poly(ADP-ribose) polymerase. *Proc. Natl. Acad. Sci. USA* **84**, 8370–8374.
- Cleaver, J. E., and Kraemer, K. H. (1989). Xeroderma pigmentosum. In *The Metabolic Basis of Inherited Disease*, Vol. II, C. R. Scriver, A. L. Beaudet, W. S. Sly, and D. Valle, eds. (New York: McGraw-Hill), pp. 2949–2971.
- Collins, A., and Johnson, R. T. (1987). DNA repair mutants in higher eukaryotes. *J. Cell Sci. (Suppl.)* **6**, 61–82.
- de Jonge, A. J. R., Vermeulen, W., Klein, B., and Hoeijmakers, J. H. J. (1983). Microinjection of human cell extracts corrects xeroderma pigmentosum defect. *EMBO J.* **2**, 637–641.
- de Vos, A. M., Tong, L. M., Milburn, M. V., Natias, P. M., Jancarik, J., Noguchi, S., Nishimura, S., Miura, K., Ohtsuka, E., and Kim, S. H. (1988). Three dimensional structure of an oncogene protein: catalytic domain of human c-H-ras p21. *Science* **239**, 888–893.
- Doolittle, R. F., Johnson, M. S., Husain, I., van Houten, B., Thomas, D. C., and Sancar, A. (1986). Domain evolution of a prokaryotic DNA repair protein and its relationship to active-transport proteins. *Nature* **323**, 451–453.
- Ford, M. J., Anton, I. A., and Lane, D. P. (1988). Nuclear protein with sequence homology to translation initiation factor eIF-4A. *Nature* **332**, 736–738.
- Friedberg, E. C. (1985). *DNA Repair* (San Francisco: W. H. Freeman and Co.).
- Friedberg, E. C. (1988). DNA repair in the yeast *Saccharomyces cerevisiae*. *Microbiol. Rev.* **52**, 70–102.
- Fu, X.-Y., Ge, H., and Manley, J. L. (1988). The role of polypyrimidine stretch at the SV40 early pre-mRNA 3' splice site in alternative splicing. *EMBO J.* **7**, 809–817.
- Gorbalenya, E. G., Koonin, E. V., Donchenko, A. P., and Blinov, V. M. (1988). A conserved NTP-motif in putative helicases. *Nature* **333**, 22.
- Gorbalenya, E. G., Koonin, E. V., Donchenko, A. P., and Blinov, V. M. (1989). Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucl. Acids Res.* **17**, 4713–4730.
- Graham, F., and van der Eb, A. J. (1973). A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52**, 456–467.
- Grossman, L., Caron, P. R., Mazur, S. J., and Oh, E. Y. (1988). Repair of DNA containing pyrimidine dimers. *FASEB J.* **2**, 2696–2701.
- Hay, B., Jan, L. Y., and Jan, Y. N. (1988). A protein component of *Drosophila* polar granules is encoded by *vasa* and has extensive sequence similarity to ATP-dependent helicases. *Cell* **55**, 577–587.
- Haynes, R. H., and Kunz, B. A. (1981). DNA repair and mutagenesis in yeast. In *The Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance*, J. Strathern, E. Jones, and J. Broach, eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory), pp. 371–414.
- Higgins, D. R., Prakash, S., Reynolds, P., Polakowska, R., Weber, S., and Prakash, L. (1983). Isolation and characterization of the *RAD3* gene of *Saccharomyces cerevisiae* and inactivation of *RAD3* deletion mutants. *Proc. Natl. Acad. Sci. USA* **80**, 5680–5684.
- Hodgman, T. C. (1988). A new superfamily of replicative proteins. *Nature* **333**, 23–23, 578 [Erratum].
- Hoeijmakers, J. H. J., van Duin, M., Westerveld, A., Yasui, A., and Bootsma, D. (1986). Identification of DNA repair genes in the human genome. *Cold Spring Harbor Symp. Quant. Biol.* **51**, 91–101.
- Hoeijmakers, J. H. J., Odijk, H., and Westerveld, A. (1987). Differences between rodent and human cell lines in the amount of integrated DNA after transfection. *Exp. Cell Res.* **169**, 111–119.
- Jentsch, S., McGrath, J. P., and Varshavsky, A. (1987). The yeast DNA repair gene *RAD6* encodes a ubiquitin-conjugating enzyme. *Nature* **329**, 131–134.
- Johnson, R. T. (1989). Reply to letter by J. H. Robbins. *Hum. Genet.* **84**, 101.
- Johnson, R. T., Ellito, G. C., Squires, S., and Joysey, V. C. (1989). Lack of complementation between xeroderma pigmentosum complementation groups D and H. *Hum. Genet.* **81**, 203–210.
- Jones, J. S., Weber, S., and Prakash, L. (1988). The *Saccharomyces cerevisiae RAD18* gene encodes a protein that contains potential zinc finger domains for nucleic acid binding and a putative nucleotide binding sequence. *Nucl. Acids Res.* **14**, 7119–7131.
- Kodahek, T., and Alberts, B. (1987). Stimulation of protein strand exchange by a DNA helicase. *Nature* **326**, 312–314.
- Kozak, M. (1984). Compilation and analysis of sequences upstream from the translation start site in eukaryotic mRNAs. *Nucl. Acids Res.* **12**, 857–872.
- Kozak, M. (1989). The scanning model for translation: an update. *J. Cell Biol.* **108**, 229–241.
- Lapeyre, B., Bourbon, H., and Amalric, F. (1987). Nucleolin, the major protein of growing eukaryotic cells: an unusual protein structure revealed by the nucleotide sequence. *Proc. Natl. Acad. Sci. USA* **84**, 1472–1476.
- Lehmann, A. R. (1987). Cockayne's syndrome and trichothiodystrophy: defective repair without cancer. *Cancer Rev.* **7**, 82–103.
- Linder, P., Lasko, P. F., Nielsen, P. J., Nishi, K., and Slonimski, P. P. (1989). Birth of the D-E-A-D box. *Nature* **337**, 121–122.
- Madura, K., and Prakash, S. (1986). Nucleotide sequence, transcript mapping and regulation of the *RAD2* gene of *Saccharomyces cerevisiae*. *J. Bacteriol.* **166**, 914–923.
- Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

- Matson, S. W. (1986). *Escherichia coli* helicase II (*uvrD* gene product) translocates unidirectionally in a 3' to 5' direction. *J. Biol. Chem.* 261, 10169–10175.
- Metherall, J. E., Collins, F. S., Pan, J., Weissman, S. M., and Forget, G. (1986). β^0 thalassemia caused by a base substitution that creates an alternative splice acceptor site in an intron. *EMBO J.* 5, 2551–2557.
- Nakano, T., and Suzuki, K. (1989). Genetic cause of a juvenile form of Sandhoff disease. *J. Biol. Chem.* 264, 5155–5158.
- Naumovski, L., and Friedberg, E. C. (1983). A DNA repair gene required for the incision of damaged DNA is essential for viability in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 80, 4818–4821.
- Nishi, K., Morel-Deville, F., Hershey, J. W. B., Leighton, T., and Schier, J. (1988). An *elf-4A*-like protein is a suppressor of an *Escherichia coli* mutant defective in 50S ribosomal subunit assembly. *Nature* 336, 496–499.
- Oh, E. Y., and Grossman, L. (1987). Helicase properties of the *Escherichia coli* UvrAB complex. *Proc. Natl. Acad. Sci. USA* 84, 3638–3642.
- Okayama, H., and Berg, P. (1982). A cDNA cloning vector that permits expression of cDNA inserts in mammalian cells. *Mol. Cell. Biol.* 78, 4818–4821.
- Pabo, C. O., and Sauer, R. T. (1984). Protein–DNA recognition. *Annu. Rev. Biochem.* 53, 293–321.
- Ptashne, M. (1988). How eukaryotic transcriptional activators work. *Nature* 335, 683–689.
- Reeck, G. R., Isackson, P. J., and Teller, D. C. (1982). Domain structure in high molecular weight high mobility group nonhistone chromatin proteins. *Nature* 300, 76–78.
- Robbins, J. H. (1989). No lack of complementation for unscheduled DNA synthesis between xeroderma pigmentosum complementation groups D and H. *Hum. Genet.* 84, 99–100.
- Roberts, B. (1989). Nuclear location signal-mediated protein transport. *Biochim. Biophys. Acta* 1008, 263–280.
- Saiki, R., Scharf, S., Faloona, F., Mullis, K., Horn, G., Ehrlich, H. A., and Arnheim, N. (1985). Enzymatic amplification of β -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230, 1350–1353.
- Sancar, A., and Sancar, G. B. (1988). DNA repair enzymes. *Annu. Rev. Biochem.* 57, 29–67.
- Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.
- S  raphin, B., Simon, M., Boulet, A., and Faye, G. (1989). Mitochondrial splicing requires a protein from a novel helicase family. *Nature* 337, 84–87.
- Shaw, G., and Kamen, R. (1986). A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. *Cell* 46, 659–667.
- Shepherd, J. C. W., McGinnis, W., Carrasco, A. E., De Robertis, E. M., and Gehring, W. J. (1984). Fly and frog homeo domains show homologies with yeast mating type regulatory proteins. *Nature* 310, 70–71.
- Smith, C. W. J., Porro, E. B., Patton, G., and Nadal-Ginard, B. (1989). Scanning from an independently specified branch point defines the 3' splice site of mammalian introns. *Nature* 342, 243–247.
- Sopata, M., Burton, Z. F., and Greenblatt, J. (1989). Structure and associated DNA-helicase activity of a general transcription initiation factor that binds to RNA polymerase II. *Nature* 341, 410–414.
- Spritz, R. A., Jagadeeswaran, P., Choudary, P. V., Biro, P. A., Elder, J. T., deRiel, J. K., Manley, J. L., Geffer, M. L., Forget, B. G., and Weissman, S. M. (1981). Base substitution in an intervening sequence of a β^+ thalassemic human globin gene. *Proc. Natl. Acad. Sci. USA* 78, 2455–2459.
- Staden, R., and McLachlan, A. D. (1982). Codon preference and its use in identifying protein coding regions in long DNA sequences. *Nucl. Acids Res.* 10, 141–156.
- Stefanini, M., Keijzer, W., Westerveld, A., and Bootsma, D. (1985). Interspecies complementation analysis between xeroderma pigmentosum and UV-sensitive Chinese hamster cells. *Exp. Cell Res.* 161, 373–380.
- Struhl, K. (1989). Helix-turn-helix, zinc-finger, and leucine-zipper motifs for eukaryotic transcriptional regulatory proteins. *Trends Biochem. Sci.* 14, 137–140.
- Sung, P., Prakash, L., Matson, S. W., and Prakash, S. (1987a). RAD3 protein of *Saccharomyces cerevisiae* is a DNA helicase. *Proc. Natl. Acad. Sci. USA* 84, 8951–8955.
- Sung, P., Prakash, L., Weber, S., and Prakash, S. (1987b). The RAD3 gene of *Saccharomyces cerevisiae* encodes a DNA-dependent ATPase. *Proc. Natl. Acad. Sci. USA* 84, 6045–6049.
- Tanaka, K., Satokata, I., Ogita, Z., Uchida, T., and Okada, Y. (1989). Molecular cloning of a mouse DNA repair gene that complements the defect of group-A xeroderma pigmentosum. *Proc. Natl. Acad. Sci. USA* 86, 5512–5516.
- Teitz, T., Naiman, T., Avissar, S. S., Bar, S., Okayama, H., and Canaan, D. (1987). Complementation of the UV-sensitive phenotype of a xeroderma pigmentosum human cell line by transfection with a cDNA clone library. *Proc. Natl. Acad. Sci. USA* 84, 8801–8804.
- Thompson, L. H. (1989). The somatic cell genetics approach to dissecting mammalian DNA repair. *Environ. Mol. Mutagen.* 14, 264–281.
- Thompson, L. H., Rubin, J. S., Cleaver, J. E., Whitmore, G. F., and Brookman, K. W. (1980). A screening method for isolating DNA repair-deficient mutants of CHO cells. *Somatic Cell Genet.* 6, 391–405.
- Thompson, L. H., Busch, D. B., Brookman, K. W., Mooney, C. L., and Glaser, D. A. (1982). Genetic diversity of UV-sensitive DNA repair mutants of Chinese hamster ovary cells. *Proc. Natl. Acad. Sci. USA* 78, 3734–3737.
- Thompson, L. H., Mooney, C. L., and Brookman, K. W. (1985). Genetic complementation between UV-sensitive CHO mutants and xeroderma pigmentosum fibroblasts. *Mutat. Res.* 150, 423–429.
- Thompson, L. H., Shiomi, T., Salazar, E. P., and Stewart, S. A. (1988). An eighth complementation group of rodent cells hypersensitive to ultraviolet radiation. *Somatic Cell Mol. Genet.* 14, 605–612.
- Uemura, T., Morino, K., Uzawa, S., Shiozaki, K., and Yanagida, M. (1987). Cloning and sequencing of *Schizosaccharomyces pombe* DNA topoisomerase I gene, and effect of gene disruption. *Nucl. Acids Res.* 15, 9727–9739.
- van Duin, M., de Wit, J., Odijk, H., Westerveld, A., Yasui, A., Koken, M. H. M., Hoeijmakers, J. H. J., and Bootsma, D. (1986). Molecular characterization of the human excision repair gene *ERCC-1*: cDNA cloning and amino acid homology with the yeast DNA repair gene *RAD10*. *Cell* 44, 913–923.
- van Duin, M., Jansen, J. H., de Wit, J., Hoeijmakers, J. H. J., Thompson, L. H., and Bootsma, D. (1988). Transfection of the cloned human excision repair gene *ERCC-1* to UV-sensitive CHO mutants only corrects the repair defect in complementation group-2 mutants. *Mutat. Res.* 193, 123–130.
- van Duin, M., Vredeveltd, G., Mayne, L. V., Odijk, H., Vermeulen, W., Klein, B., Weeda, G., Hoeijmakers, J. H. J., and Bootsma, D. (1989). The cloned human excision repair gene *ERCC-1* fails to correct xeroderma pigmentosum complementation group A through I. *Mutat. Res.* 217, 83–92.
- Venema, J., van Hoffen, A., Natarajan, A. T., van Zeeland, A. A., and Mullenders, L. H. F. (1990a). The residual repair capacity of xeroderma pigmentosum complementation group C fibroblasts is highly specific for transcriptionally active DNA. *Nucl. Acids Res.* 18, 443–448.
- Venema, J., Mullenders, L. H. F., Natarajan, A. T., van Zeeland, A. A., and Mayne, L. (1990b). The genetic defect in Cockayne's syndrome is associated with a defect in repair of UV-induced DNA damage in transcriptionally active DNA. *Proc. Natl. Acad. Sci. USA*, in press.
- Verlaan-de Vries, M., Bogaard, M. E., van de Elst, H., van Boom, H. J., van der Eb, A. J., and Bos, J. L. (1986). A dot-blot screening procedure for mutated *ras* oncogenes using synthetic oligodeoxynucleotides. *Gene* 50, 313–320.
- Vermeulen, W., Ossewey, P., de Jonge, A. J. R., and Hoeijmakers, J. H. J. (1986). Transient correction of excision repair defects in fibroblasts of 9 xeroderma pigmentosum complementation groups by microinjection of crude human cell extracts. *Mutat. Res.* 165, 199–206.
- Walker, J. E., Saraste, M., Runswick, J. J., and Gay, N. J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP syn-

thase, myosin, kinases and other ATP-requiring enzymes. *EMBO J.* 1, 945–951.

Weber, C. A., Salazar, E. P., Stewart, S. A., and Thompson, L. H. (1988). Molecular cloning and biological characterization of a human gene, *ERCC-2*, which corrects the nucleotide excision repair defect in CHO UV5 cells. *Mol. Cell. Biol.* 8, 1137–1146.

Weber, C. A., Salazar, E. P., Stewart, S. A., and Thompson, L. H. (1990). *ERCC-2*: cDNA cloning and molecular characterization of a human nucleotide excision repair gene with high homology to yeast *RAD3*. *EMBO J.* 9, 1437–1447.

Weeda, G., van Ham, R. C. A., Masurel, R., Westerveld, A., Odijk, H., de Wit, J., Bootsma, D., van der Eb, A. J., and Hoeijmakers, J. H. J. (1990). Molecular cloning and biological characterization of the human excision repair gene *ERCC-3*. *Mol. Cell. Biol.* 10, 2570–2581.

Wen, L., Huang, J., Johnson, B. H., and Reeck, G. R. (1989). A human placental cDNA clone that encodes nonhistone chromosomal protein HMG-1. *Nucl. Acids Res.* 17, 1197–1214.

Westerveld, A., Hoeijmakers, J. H. J., van Duin, M., de Wit, J., Odijk, H., Pastink, A., Wood, R. D., and Bootsma, D. (1984). Molecular cloning of a human DNA repair gene. *Nature* 310, 425–428.

Wickens, M., and Stephenson, P. (1984). Role of a conserved AAUAAA sequence: four AAUAAA point mutations prevent messenger 3' end formation. *Science* 226, 1045–1051.

Wood, R. D., and Burki, H. J. (1982). Repair capability and the cellular age response for killing and mutation induction after UV. *Mutat. Res.* 95, 505–514.

Zdzienicka, M. Z., and Simons, J. W. I. M. (1986). Analysis of repair processes by the determination of the induction of cell killing and mutations in two repair-deficient Chinese hamster ovary cell lines. *Mutat. Res.* 166, 59–69.

Zdzienicka, M. Z., and Simons, J. W. I. M. (1987). Mutagen-sensitive cell lines are obtained with a high frequency in V79 Chinese hamster cells. *Mutat. Res.* 178, 235–244.

Zdzienicka, M. Z., van der Schans, G. P., Westerveld, A., van Zeeland, A. A., and Simons, J. W. I. M. (1987). Phenotypic heterogeneity within the first complementation group of UV-sensitive mutants of Chinese hamster cell lines. *Mutat. Res.* 193, 31–41.

GenBank Accession Number

The accession number for the sequence reported in this paper is M31899.